VERY PRELIMINARY AND INCOMPLETE: DO NOT CITE

The Evolution of National Retail Chains: How We Got Here*

Lucia Foster
Bureau of the Census
Lucia.S.Foster@census.gov

John Haltiwanger
University of Maryland
and NBER
haltiwan@econ.umd.edu

C.J. Krizan
Bureau of the Census
Cornell.J.Krizan@census.gov

September 2006

* This paper was written in part by Census Bureau Staff. It has undergone a more limited review than official Census Bureau publications. All results were reviewed to ensure confidentiality. Any views, findings and opinions in the paper reflect the views of the authors and do not reflect the views of the U.S. Census Bureau. We thank Ron Jarmin and Javier Miranda for sharing the Longitudinal Business Database and the bridge between the Longitudinal Business Database and COMPUSTAT with us. We thank Kristin McCue for sharing her 2002 firm identifier.

Abstract

The growth and dominance of large national chains is a ubiquitous feature of the U.S. retail sector. The increasing dominance of these chains, and their impact on the size distribution and firm turnover rates in retail trade, has been documented by Jarmin et al. (2005). Moreover, these large, national chains drove most of the recent productivity growth in the U.S. retail trade industry as more productive entering establishments affiliated with national chains displaced much less productive exiting single-unit firms (Foster et al (forthcoming)). An open question to explore concerns the factors that prompted these dynamics and in turn the factors that led some national chains like Wal-Mart, Starbucks and Olive Garden to succeed dramatically relative to other national chains. Holmes (2005) explores some of the issues in exploring the dynamics of the location of Wal-Mart establishments in the U.S. We build on this literature by following the paths of large, national chain retail firms from 1977 to 2002 using establishment and firm-level data from the Census of Retail Trade and the Longitudinal Business Database. We begin by exploring the paths to success of the dominant national chains in 2002. In particular, we look for patterns in the geographic expansion, size, and transition from privately-held to publicly-owned ownership for the firms that came to dominate the retail trade industry over this twenty-five year period and ask how they differ from those that failed.

1. Introduction

A ubiquitous feature of the U.S. retail industry is the growth and dominance of large, national chains in the retail industry. The best known is Wal-Mart but large, national chains in many different retail sectors have become paramount. Jarmin et al. (2005) document the increasing preeminence of large, national chains and their impact on the size distribution and firm turnover rates in retail trade. Foster et al. (forthcoming) show that virtually all of the productivity growth in the U.S. retail trade market over the 1990s is due to more productive entering establishments affiliated with large, national chains displacing much less productive exiting establishments that are "mom and pop" single-unit establishment firms. An open question to explore concerns the factors that prompted these dynamics and in turn the factors that led some national chains like Wal-Mart, Starbucks and Olive Garden to succeed dramatically relative to other national chains. Some existing studies have started to investigate these patterns. For example, Holmes (2005) explores some of the issues in exploring the dynamics of the location of Wal-Mart establishments in the U.S.

We build on this literature by following the paths of large, national chain retail firms from 1977 to 2002 using establishment and firm-level data from the Census of Retail Trade and the Longitudinal Business Database. We begin by exploring the paths to success of the preeminent national chains in 2002. In particular, we look for patterns in the geographic expansion, size, and transition from privately-held to publicly-owned, for the firms that came to dominate the retail trade industry over this twenty-five year period and ask how they differ from those that failed.

Not all retail industries are alike, there is likely to be a range of many chains' activity across industries. Furthermore, it is clear that a company such as Wal-Mart uses a very different

business model than does McDonald's. Besides the obvious differences in establishment size, there may also be important differences in geographic expansion patterns, their employment growth rates, the number of establishments per firm and the propensity of being a publicly traded firm. An important aspect of our analysis is to document differences in patterns across firm types and industries in the path towards becoming a large, national firm.

The paper proceeds as follows. Section 2 provides a brief literature review to help put the questions and approach of the paper into context. Section 3 describes the data used in the analysis. Section 4 presents an overview of the patterns of structural change in the retail trade sector. This section very much builds on the recent literature and it is clear from this section as well as the recent literature that large, national chain retail firms increasingly dominate the retail trade sector. Section 5 begins the exploration of the evolution of the large, national firms that became paramount in the retail trade sector by 2002. Concluding remarks are given in section 6.

2. Literature Review

As noted in the introduction, the dramatic changes in the retail trade sector have yielded a burgeoning literature documenting and exploring the factors underlying the changes. Jarmin et. al. (2004 and 2005) take advantage of the newly developed Longitudinal Business Database (LBD) to document and analyze the patterns of growth and change in the retail trade sector. They find an increasingly dominant role of large, national chains in retail trade activity as measured by payroll and employment. Moreover, they quantify the extent and patterns of firm and establishment entry and exit in the retail trade sector. Not surprisingly, entry and exit rates are relatively high in retail trade (e.g., compared to manufacturing) but interestingly they find that entering establishments are much larger in terms of relative size to incumbents compared to

their manufacturing counterparts. They find that the patterns of activity and change vary across market size and type. Rural areas are still served by a relatively large number of single unit establishment firms but even in rural areas such firms are experiencing net losses. In larger markets, there is higher firm turnover. They also find that single unit firms and large, national chains are more likely to coexist in some industries (such as Eating and Drinking) and less likely in others like General Merchandise stores.

One outstanding example of these dramatic changes in retail trade is Wal-Mart. Not surprisingly, some of the literature documenting and studying the changing structure of the retail trade sector focuses on Wal-Mart. Basker (2005) studies the labor market effects of Wal-Mart entry into a local market. Basker finds that there is an initial increase in employment in the local area but over the subsequent five years a Wal-Mart entry yields exits and contractions by competitors. She also finds evidence for upstream effects via a decline in wholesalers employment. Holmes (2006) also explores Wal-Mart entry dynamics but with a different perspective. Holmes documents the geographic pattern of expansion of Wal-Mart. Wal-Mart started in Bentonville, Arkansas and its expansion path shows that it first expanded into local, then regional and then finally national markets. Holmes models this problem as a tradeoff between taking advantage of the economies of density (favors operations in close proximity) versus locating in the market with the highest quality (unlikely to favor operations in close proximity).

These dramatic changes in the structure of the retail trade sector have been associated with the productivity growth in the retail trade sector over this period of time. Foster et. al. (forthcoming) show that virtually all of the labor productivity growth in the U.S. retail trade sector over the 1990s is accounted for by more productive entering establishments displacing much less productive exiting establishments. Interestingly, the large productivity gap between

low productivity exiting *single-unit* establishments and entering high productivity establishments from large, *national chains* plays a disproportionate role in these dynamics.

While much has been learned from this burgeoning literature, our understanding of these structural changes in the retail trade sector is still limited. The recent development of the LBD (see Jarmin and Miranda, 2002) has provided a rich new resource in its own right for the study of these issues (as is evident in Jarmin et. al. (2004, 2005)) but also greatly enhances the ability to use the Census of Retail Trade at the establishment and firm level for an extended period of time. The LBD provides the longitudinal establishment and firm identifiers to conduct longitudinal analysis of firms over an extended period of time (1975 to the present). The current paper takes advantage of the integration of the LBD with the Census of Retail Trade (CRT) from 1977 to 2002 to expand our understanding of the major structural changes ongoing in the retail trade sector.

3. Data and Measurement Issues

The empirical analysis in this paper uses data from the Census of Retail Trade (CRT). The Census Bureau conducts a survey of retail trade establishments every five years (those years ending in '2' and '7'). The survey questionnaire is mailed out to all large and medium-sized firms and generally all firms that operate multiple establishments; most very small firms are excused from answering the questionnaire. The data for these very small firms come from two sources: either a Census sample of these very small firms or administrative records from other federal agencies. We use both reported data and administrative data in our empirical exercises because there is no reason to suppose that the administrative records data are inferior to the reported data for the variables being used in this study. The CRT contains data on

establishments concerning the kind of business, physical location, sales in dollars, annual and first quarter payroll, and employment for the pay period including March 12th. We create a measure of real sales by deflating nominal sales by the Consumer Price Index in each year (in 2002 dollars).[1] High quality establishment-level data are available from the CRT for the period 1977-2002. During this period, the industry classification system used by Census in collecting data switched from the Standard Industrial Classification (SIC) to North American Industry Classification System (NAICS). In order to ensure comparability across time, we define the retail trade sector using the SIC definition for our entire sample period. Fortunately, the Census Bureau maintained SIC industry information in the NAICS transition year (1997) so creating the 1997 retail trade sector under SIC is straightforward. It is slightly more difficult to create the 2002 retail trade sector under SIC. We have created a rough version of 2002 retail trade sector under SIC by applying 1997 SIC codes to continuing establishments in 2002 and using a SIC-NAICS concordance to translate the industry codes for establishment births in 2002 (see the Data Appendix for more details). To the extent that we cannot remove all establishments that first appear in the 2002 CRT and are in NAICS-only industries, our sample (slightly) overestimates the 2002 retail trade sector on an SIC consistent basis.

The retail trade sector grows significantly during our sample period, but roughly speaking, there are about 1.5 million establishments with 20 million paid employees generating $2.5 trillion in real sales in retail trade in a given year. The CRT also collects information on firm ownership of establishments. There are approximately 1 million firms in retail trade in each year in our sample. It is apparent from the relative magnitudes of the number of firms and establishments that most firms in retail trade are single-unit establishments.

[1] In future drafts of this paper we hope to use the four-digit industry price deflator from BLS instead of the Consumer Price Index.

We restrict our sample to establishments that can be matched to the LBD because we will use LBD data in part of our analysis. We further restrict our analysis to establishments that have positive employment (and sales and payroll). This restriction on positive employment helps us to increase the quality of our data in early Census years. Our restrictions cause us to remove approximately 100,000 establishments. Our sample exhibits qualitatively similar time series patterns as the published data of the full universe of the retail trade sector.

We define chains in terms of the number of states in which a firm has establishments operating (or the number of establishments in operation for single-state firms). We classify firms into five chain types. Firms that have a single establishment are classified as Single. Firms that have multiple establishments operating in a single state are classified as Local. Firms that have multiple establishments operating over 2 to 5 states are classified as Regional. Firms that have multiple establishments operating over 6 to 14 states are classified as National. Finally, firms that have multiple establishments operating over 15 or more states are classified as Megas. The cutoffs of these different types of chains in terms of number of states are arbitrary, but as will become clear, are quite instructive for characterizing the changing structure of the U.S. retail trade industry.

In some of our analysis, we consider whether the firm is publicly traded. Our data on whether the firm is publicly traded are derived from merging the LBD with COMPUSTAT. Unfortunately, we do not yet have this data for 2002.

4. Trends in Chain Types

Figure 1 presents the total number of firms and establishments (upper panel) and total real sales and employment (lower panel) for our sample of the retail trade sector from 1977 to

2002. The number of firms is large (about 1 million) and relatively constant over this time period. The number of establishments, not surprisingly, is much larger and the gap between the number of firms and establishments has risen steadily. By 2002 there are approximately 1.5 million retail trade establishments. The number of employees and real sales is also very large. In 2002, the retail trade sector accounted for about 25 million workers and over $3 trillion in gross real sales (2002 dollars). Employment and sales both grew rapidly over the period.

Figure 2 presents total employment (top panel) and total real sales (lower panel) for single unit firms (SU) and multiple establishment firms (MU) for the retail trade sector from 1977 to 2002. Interestingly, in 1977 employment and sales for single units and multiunits were about the same but since then employment and sales for multiunit firms have increased rapidly relative to that of SUs. By 2002, MUs account for roughly two-thirds of sales and employment.

Figure 3 presents total employment (top panel) and total real sales (lower panel) by chain type and it shows dramatic growth in employment and sales for Mega retail firms from 1977 to 2002. It is apparent that the growth in multiunit sales and employment in Figure 2 is dominated by the growth in sales and employment for Mega firms. There is some modest growth in employment and sales for the other multiunit chain types but this is dwarfed by the changes in Mega firms. The employment shares for the retail trade by chain type show little change for Local, Regional and National chains, a dramatic decrease in for Singles (from 50% to 38%) and a concomitant dramatic increase for Megas (from 22% to 35%).

Figure 4 shows the total number of firms (in logs) by chain type. The total number of firms in 2002 is about the same as in 1977. Single firms dominate in terms of the number of firms over the entire sample period, but there has been substantial growth in the number of Mega firms. The number of Mega firms more than doubled over this period. There was also substantial growth in the number of Regional and National firms. Interestingly, the number of

Local and Regional firms peaked in 1987 (for Local the number has been declining since then). Figure 5 shows the analogous total number of establishments (in logs) by chain type. The total number of establishments associated with Mega firms more than doubles over the period. There is substantial growth in the number of establishments associated with Regional and National firms, although the number of National firms' establishments has been declining since 1987. The number of establishments associated with Local firms also peaks in 1987 and then declines thereafter so that it is nearly the same in 2002 and 1977.

Figure 6 shows the average of the log number of establishments per firm by chain type. Not surprisingly, the number of establishments per firm is much larger for Mega firms than for other firms. The difference between the number of establishments per firm by Mega firms and National firms is over about 200 log points. Recalling the results from earlier figures, we find that while the total number of establishments has been relatively constant over the 1977 to 2002 period, the composition of establishments has changed substantially with a dramatic increase in the number of establishments from Mega firms. Interestingly, the latter is largely driven by an increase in the number of Mega firms rather than in the number of establishments per Mega firm.

Figure 7 shows the average employment per firm (in logs) and Figure 8 shows the average real sales per firm (in logs) by chain type. Not surprisingly, Mega firms are much larger than other firms in terms of sales and employment. Megas are more than 700 log points larger in employment and sales than SUs. There is some modest growth in average employment size for all chain types and relatively flat average sales per firm. Returning to Figures 1-3, it is clear that the rapid increase in sales and employment in the retail trade sector is associated with an increase in the role of Mega firms and in particular the number of Mega firms.

Further information about the size distribution of employment across chain types is

presented in Table 1. The top panel shows average employment size of the different types of firms. In Table 1 we see stark differences between the average size of single establishment firms and Megas. However, even these stark average size differences do not fully capture the drastic differences in the size distribution. The lower panel of Table 1 shows the average size of firms weighted by employment. This statistic provides a summary measure of the coworker mean which is the size of the average firm for the average worker (see Davis, Haltiwanger and Schuh (1996)). Computing these weighted statistics is interesting because of the skewness of the firms' size distribution.

The bottom panel shows that in 2002 the average employee in a Single firm had 66 co-workers, a significant increase since 1977. However Mega firm employees had far more co-workers on average: 184,369. Compare this to the simple average Mega firm size of 14,911. The difference between the coworker mean and the simple mean provides information about the skewness of the size distribution. An interesting feature of Table 1 is that the distribution of Mega firm sizes is very skewed in its own right. Put differently, while Mega *firms* are large on average, their size distribution is very skewed so that the average *worker* at a Mega firm works for a very large firm.

The main point of this introductory analysis is the increasingly dominant role of Mega firms in retail trade. In our subsequent analysis, we seek to explore the nature of the dramatic increase in the role of Mega firms. We turn to that analysis in the next section.

5. Mega Firms

Given the increased dominance of Mega firms in retail trade, it is clearly of interest to understand the path and the factors that yielded this phenomenon. In this section, we explore the dynamics of the Mega firms by focusing on the Mega firms in 2002 and asking how they evolved over the 25 years of our sample. To begin this analysis, we examine the dynamics of these firms by the year of entry of the firm.

Table 2 shows the transition dynamics of the 2002 Mega firms by birth year of firm and the tenure of the firm as a Mega firm. For example, a firm born in 1987 that became a Mega in 1992 is part of the birth year cohort of 1987 and tenure group of 10 (since by 2002 it had been a Mega for 10 years). The top panel shows the transition matrix in terms of employment, the bottom in real sales. The first column of each panel shows the share of 2002 activity accounted for by the overall birth cohort. The next six columns show the contribution from each tenure cohort (where the sum of columns 2 through 6 is equal by construction to the total share in column 1). There is left censoring of the birth cohorts and right censoring of the tenure. Given the censoring, some of the elements in columns 2 through 6 are omitted by construction. For example, there are no firms with tenure equal to 25 for birth cohort 1997. Additionally, some cells are suppressed for disclosure reasons (denoted by D).

Most of the activity in 2002 of the Mega firms is from firms that existed in 1977. Nonetheless, about 30 percent of activity is from later cohorts. Interestingly, even though the firms in existence in 1977 predominate, only about half of those firms were Megas in 1977. Across all birth cohorts, about 15 percent of sales and employment of Megas in 2002 are associated with Megas with tenure less than 5 years old. Almost half of the activity of Megas in 2002 is associated with firms with tenure as Megas less than 15 years old. Taking the birth

cohorts and the tenure as Megas together, a significant fraction of the mega activity in 2002 is due to new mega firms, either true new entrants or firms that have become mega in the last 15 years.

Given that the 2002 Mega firms represent a rich blend of older and newer firms, we explore a cohort analysis of the 2002 Mega firms. For this analysis, we focus on the cohort analysis by birth year of the firm and restrict our analysis to the 2002 Mega firms. By doing this we intentionally introduce sample selection since we want to illuminate the evolution of the 2002 firms (i.e., how we got here). In conducting this analysis, we are not identifying when the Mega 2002 firm became a Mega but rather when it first started operation as any type of firm. As will become clear, many new cohorts of firms that are ultimately Mega firms by 2002 are not Mega firms initially and it takes time for these firms to become Mega firms.

When interpreting this cohort analysis several cautions should be kept in mind. First, as noted earlier, there are issues of left censoring for the 1977 birth cohort and right censoring for all cohorts. Second, when we define firm birth as the entry of a firm (legal entity) as identified by a new firm identifier, this "firm birth" may in fact reflect change in ownership structure (via , e.g., merger and acquisition activity) as well as de novo firms. While other studies (Davis et. al. (forthcoming)) show that the predominant number of new firms in the LBD are truly new firms (with only new establishments), it is less clear that this is the case for very large Mega firms. Put differently, most firms are single unit firms so the distinction between firm and establishment age is not important – for mega firms that have many establishments, this distinction is likely important.

Since a firm is a legal entity that owns physical establishments, it is possible to define the first year of the firm using either the first appearance of the legal entity or of the oldest physical establishment that the firm owns. We examine this distinction in the cross tabulation

shown in Table 3. Table 3 shows firm first years for the 2002 Megas defined by the year the firm identifier first appears in the LBD (rows) and by the year the oldest establishment that the firm owns first appears (columns). It is clear that many Megas own establishments older than the firm itself. For example, 37 of the Megas whose firm identifier first appeared in 1982 (second row) had oldest establishments that first appeared in 1977 (first column). Given the size and geographic scope of these firms, perhaps this is not too surprising. What is a bit more surprising is that some of first years dated by the "oldest" establishments are younger than the parent firm. For example, ten of the Megas whose firm identifier first appeared in 1977 (first row) had oldest establishments that first appeared in 1982 (second column). To understand how this can happen, recall that all of the Table 3 results are for the 2002 Megas. It may be that a 2002 Mega firm that started up in 1977 has by 2002 divested itself or closed down all of the establishments that date back to 1977 – so the firm is older than the age of its oldest establishment. As a further note of caution, this pattern might reflect problems with longitudinal establishment linkages in the LBD.

It is clear, though, that defining firm first year by the age of the oldest establishment shifts the firm age distribution of the 2002 Megas to the right. That is, the 2002 Mega firms seem older when one defines age through oldest establishments rather than the firm identifier. To see how this may impact some of our earlier results, we turn to Table 4 which reproduces Table 2 using the oldest establishment as the identifier of firm births. It is still the case that most of the activity in 2002 of the Mega firms is from firms that existed in 1977 but few of them were Megas in 1977. Thus, in a broad sense our earlier findings are confirmed. In what follows, we restrict our analysis to using the firm age defined by the legal entity. We believe for current purposes this firm age is of interest since understanding the evolution to Mega status may be connected to merger and acquisition activity. In future drafts of the paper, however, we

will explore both types of firm age and their interaction since it is also the case that quantifying the evolution of Megas by the age of the oldest establishment is of interest. In addition, it is of interest to identify separately those firms that became Megas through merger and acquisition activity.

Figure 9 shows the evolution of average per firm employment (in logs) for the birth cohorts for the 2002 Mega firms. The plot shows that conditional on survival all the birth cohorts exhibit dramatic growth. The 1977 cohort has an average per firm employment growth of over 200 log points, the 1982 cohort over a shorter horizon has an average per firm employment growth of 350 log points and so on. Appropriate caution is required in interpreting the performance of recent cohorts given the right censoring (e.g., the 2002 cohort that enters appears to come in at a high average per firm employment but presumably this is a better indicator of the average size of Mega firms that immediately become Megas rather than the average size of firms that eventually become Megas after the point of entry). Figure 10 repeats the same type of analysis for average sales per firm (in logs) and the mimics the patterns in Figure 9. [2]

The cohort analysis helps provide an interpretation of the prior finding (in Figures 7 and 8) that showed an overall relatively modest change in average log employment and average log real sales for Mega firms. Figures 9 and 10 make clear that within a surviving cohort of Mega firms that there is substantial growth especially taking into account the firms that ultimately become mega firms. However, each new cohort comes in smaller than the incumbents and this suggests that a composition effect may be important to understand the overall changes in

[2] It is tempting to make inferences about labor productivity from comparing Figures 9 and 10. The caution in making such inference is that sales per worker differ substantially across industries and the analysis thus far does not hold industry composition fixed. Thus, spurious movements in sales per worker can occur in the aggregate as the industry composition changes. In future drafts of the paper, we will explore the role of labor productivity which

average size for all firms from one year to the next. That is, while there may be growth within birth cohorts, the entry of a new cohort with smaller average size at entry will tend to reduce the average size of all Mega firms in any given year. [3] A related point is that the growth of the Megas while Megas neglects the growth to becoming Megas. While these patterns hold for all of the cohorts, the 1982 cohort appears to have followed a slightly different path.

Figure 11 repeats the same type of analysis for the average number of establishments per firm (in logs). There is rapid growth within a cohort in the average number of establishments. For example the 1997 birth cohort exhibits more than a 200 log point growth in the average number of establishments.

Figures 12 and 13 turn to other attributes of Mega firms that have not been emphasized in earlier analysis. The average number of states per firm is depicted by birth cohort in Figure 12. Recall that by construction that a Mega firm operates in 15 or more states. However, this cohort analysis is based upon those firms that are Mega firms in 2002 and looks backwards to various birth cohorts. For the 1977 firm birth cohort, the average number of states in 1977 is around 12 suggesting that most of this cohort had not achieved Mega status. However, the 1977 firms average number of states is 16 in 1982, so that on average the 1977 birth cohort had achieved Mega status by 1982. The 1982 cohort has a slower track to Mega status. Upon birth the average number of states for the 1982 cohort is less than 6 and this cohort does not achieve Mega status on average until 1997. The 1987 cohort achieved Mega status, in contrast, on average in five years. These two cohorts show that the path to becoming a Mega firm is likely quite different across firms. Future work will explore the factors that underlie such differences.

---

will require a careful tracking of establishment and firm level industry which is a bit of a challenge over this sample given the switch from SIC to NAICS in 1997 and 2002.

Figure 13 takes another perspective on the evolution of the cohorts that become Mega firms by 2002. Figure 13 shows the fraction of the firms in each birth cohort that are publicly traded (this analysis is truncated in 1997 since the identification of publicly traded status is somewhat noisy within the current retail trade database underlying this analysis). For the 1977 birth cohort (which recall includes left censored births), more than 40 percent of the firms are publicly traded and this share grows to 60 percent by 1997. For the 1982 birth cohort, only about 20 percent of these ultimately Mega firms are publicly traded at birth but by 1997 more than 50 percent are. The rapidly growing 1987 birth cohort (as observed in Figures 9-12) starts with about 30 percent publicly traded and the share grows rapidly so that by 1997 almost 60 percent are publicly traded. To put these statistics into perspective, Davis et. al. (forthcoming) show that less than 1 percent of the firms in the U.S. private, non-agricultural sector are publicly traded so these mega firms are, not surprisingly, much more likely to be publicly traded than the typical firm. Still, it is striking that there is such a close correspondence in the growth patterns exhibited in Figures 9-12 and the share of publicly traded firms by birth cohort. While causality cannot be inferred it is clear that getting funding from going public is often part of the process of being and becoming one of the dominant mega retail trade firms in the U.S. economy.

6. Conclusions and Future Research

The growth of chain firms in the U.S. retail trade sector that has been documented in the literature is mostly a result of the growth of Mega firms. Mega retail trade firms (firms with establishments in at least 15 states) increasingly dominate the U.S. retail trade sector. Even though the difference in the number of firms between Singles and Megas is more than 900 log

[3] Figures 9 and 10 do not provide enough information to quantify the extent of this composition effect. The analysis in this section looks backward at the 2002 Mega firms by birth cohort. As is clear from Table 2, many of

points, Mega firms account for about the same amount of employment and sales as their much smaller single unit firm counterparts in 2002. The growth of Local, Regional and National chains has been far more modest. Moreover, the growth of the sales and employment by Mega firms is driven largely by an increase in the number of Mega firms, not in an increase in the average size of Megas.

Mega firms are very large. The average size of a Mega firm is almost 15,000 workers which compares to an average of around 10 workers for single unit firms. Moreover, even amongst the Mega firms, the size distribution is very skewed. The average worker who works for a Mega firm on average works for a firm with about 185,000 workers. This contrasts with the average worker for a single unit firm who works for a firm of around 66 workers.

The Mega firms in 2002 are predominately firms that have existed more than 25 years.[4] About 30 percent of 2002 Mega firms are less than 25 years old. However, for even the firms that have been in existence for at least 25 years, a substantial fraction of these firms became Mega firms in the last 15 years. Some birth cohorts of eventual Mega firms achieved Mega status much faster than other birth cohorts. The 1982 birth cohort of firms that are Mega by 2002 achieved Mega status on average in 1997. In contrast, the 1987 birth cohort of firms that are Mega by 2002 achieved Mega status on average by 1992.

A cohort analysis of the 2002 Mega firms shows that within a birth cohort there has been dramatic growth of the average size of firms. For example, the 1982 birth cohort of the 2002 Mega firms exhibited more than a 350 log point increase in average size between 1982 and 2002. The dramatic growth of the average size of firms (measured in terms of sales or

---

the Mega firms are not Mega firms upon entry but take a number of years after entry to attain Mega status.

[4] As discussed in section 5, all of these results for cohort analysis for firm age define age based upon the age of the legal entity. In future drafts we will explore alternative definitions of firm age and cohort analysis (e.g., defining firm age based on the age of the oldest establishment of the firm).

employment) within cohorts is reconciled with the modest overall increase in average size of Mega firms via composition effects. That is, while any given cohort has exhibited dramatic growth, each new cohort comes in small at entry and this composition effect reduces the average.

Not surprisingly, the share of Mega firms that are publicly traded is much larger than that for the typical firm. Still, a large fraction of Mega firms are not publicly traded. For example, even for the 1977 birth cohort of the 2002 Megas, about 40 percent of these firms are not publicly traded. There is a close correspondence between the within cohort growth of average sales per firm (or average employment per firm) and the growth in the cohorts share of publicly traded firms. For example, the rapidly growing 1987 birth cohort of 2002 Mega firms rapidly increased its share of firms that are public in a coincident pattern with growth. While causality is not clear, it is clear that part of the path to becoming a Mega firm is to go public.

The analysis in this preliminary version of the paper is largely descriptive. This paper as well as the closely related recent literature makes a strong case that understanding the structural changes towards Mega retail trade firms that operate on a large scale is important. Much work remains to be done even in a descriptive manner. Differences across industries in paths to becoming Mega have not been explored. Moreover, in the spirit of Holmes (2006), there must be economic factors that help account for the different patterns observed across birth cohorts, industries and firm types. Holmes suggests, for example, that there are potential tradeoffs between economies of density and expanding to the highest quality retail sites. The current analysis is a long way off from providing enough guidance to identify the economic factors that have led to the dominant role of Mega firms, but it is our hope that with this type of detailed description of the paths that this analysis can aid in such identification.

References

Basker, Emek (2005), "Job Creation or Destruction? Labor Market Effects of Wal-Mart Expansion," *The Review of Economics and Statistics*, 87(1), 174-83.

Davis, Steven, John Haltiwanger, and Scott Schuh (1996) *Job Creation and Destruction*. Cambridge MA: MIT Press.

Davis, Steven, John Haltiwanger, Ron Jarmin, Javier Miranda (forthcoming), "Volatility in Business Growth Rates: Publicly Traded versus Privately Held Firms," *NBER Macroeconomics Annual 2006*, Volume 21.

Foster, Lucia, John Haltiwanger, and C.J. Krizan (forthcoming), "Market Selection, Reallocation and Restructuring in the U.S. Retail Trade Sector in the 1990s," *The Review of Economics and Statistics*.

Haskel, Jonathan and Raffaella Sadun (2005), Entry, Exit and Labour Productivity in U.K. Retailing: Evidence from Micro Data," March.

Holmes, Thomas J. (2006), "The Diffusion of Wal-Mart and Economies of Density," May.

Jarmin, Ron S. and Javier Miranda (2002), "The Longitudinal Business Database," Center for Economic Studies Working Paper No. 02-17.

Jarmin, Ron S., Shawn Klimek, and Javier Miranda (2005), "The Role of Retail Chains: National, Regional, and Industry Results," Center for Economic Studies Working Paper No. 05-30.

Jarmin, Ron S., Shawn Klimek, and Javier Miranda (2004), "Firm Entry and Exit in the U.S. Retail Sector, 1977-1997," Center for Economic Studies Working Paper No. 04-17.

Data Appendix

Summary Statistics for the Retail Trade Sector

Table A1: Summary Statistics for the Retail Trade Industry, 1977-2002

Year	Establishments	Employment	Sales
SIC Basis			
1977	1,303,621	13,040,082	699,634,863
1982	1,330,316	14,467,813	1,039,028,742
1987	1,503,593	17,779,942	1,493,308,759
1992	1,526,215	18,407,453	1,894,880,209
1997	1,561,195	21,165,862	2,545,881,473
NAICS Basis			
1997	1,118,447	13,991,103	2,460,886,012
2002	1,114,637	14,647,675	3,056,421,997

Source: Census of Retail Trade, various years. The 1997 SIC data are from the Census website.

Industry Coding Issues

The 1977 and 1982 CRT have some establishments with industry codes that fall outside of retail trade. When an industry code falls outside of retail trade in our sample, we look first in the LBD for the same year and then in the next CRT. These matches greatly improve the industry codes for 1977 and 1982. Nevertheless, there are still about 200,000 in employees in 1977 and 100,000 in 1982 that are coded outside of retail trade.

There are special issues concerning a SIC-based version of retail trade in 2002. The SIC definition of retail trade is more broad than the NAICS definition of retail trade (notably it includes Eating and Drinking Places). However, part of the NAICS version of retail trade is also not under the SIC definition. These NAICS-only industries are ones that had been wholesalers selling as retailers and repair shops. Ideally we would like to remove these establishments from our sample but this is not possible. While we can identify all continuers from 1997 that should be part of SIC-based CRT this only covers continuers. We still have extra establishments in our version of the 2002 SIC-based CRT due to unclassifiable births.

Table 1: Average Size of Firms Measured By Employment

Year	Single	Local	Regional	National	Mega
Unweighted					
1977	7	38	191	1621	11249
1982	8	39	190	1427	11385
1987	9	47	216	2075	12758
1992	8	49	264	2050	11548
1997	9	56	301	2143	12367
2002	10	65	332	1834	14911
Weighted by Employment					
1977	26	606	2653	12355	92204
1982	29	757	2145	11699	80613
1987	36	1235	2689	16310	95787
1992	32	964	6949	19552	108022
1997	38	514	10939	28279	159896
2002	66	567	12999	16193	184369

Table 2: Transition Matrix by Firm Age

Cohort	Birth Share	Tenure over Census Years 0	5	10	15	20	25
Employment							
1977	68.9	5.5	3.9	3.5	3.0	17.1	35.9
1982	6.6	1.4	1.4	D	D	D	
1987	7.7	0.6	3.2	2.2	1.8		
1992	4.2	0.9	1.7	1.6			
1997	6.1	1.2	4.9				
2002	6.4	6.4					
Real Sales							
1977	72.6	5.8	4.4	2.3	2.5	21.7	35.9
1982	7.5	1.2	1.0	D	D	D	
1987	7.5	0.9	4.2	1.2	1.3		
1992	3.2	1.2	0.8	1.2			
1997	5.2	2.7	2.5				
2002	3.9	3.9					

Note: Cells that cannot be disclosed are denoted by “D.”

Table 3: Comparing First Year of Firm and of Oldest Establishment at the Firm (Number of Firms)

First Year	First Year of Oldest Establishment at the Firm					
of Firm	1977	1982	1987	1992	1997	2002
1977	205	10	D	D	D	D
1982	37	13	D	D	D	D
1987	45	D	13	D	D	D
1992	41	10	D	16	D	D
1997	34	6	9	7	7	D
2002	37	D	8	9	D	9

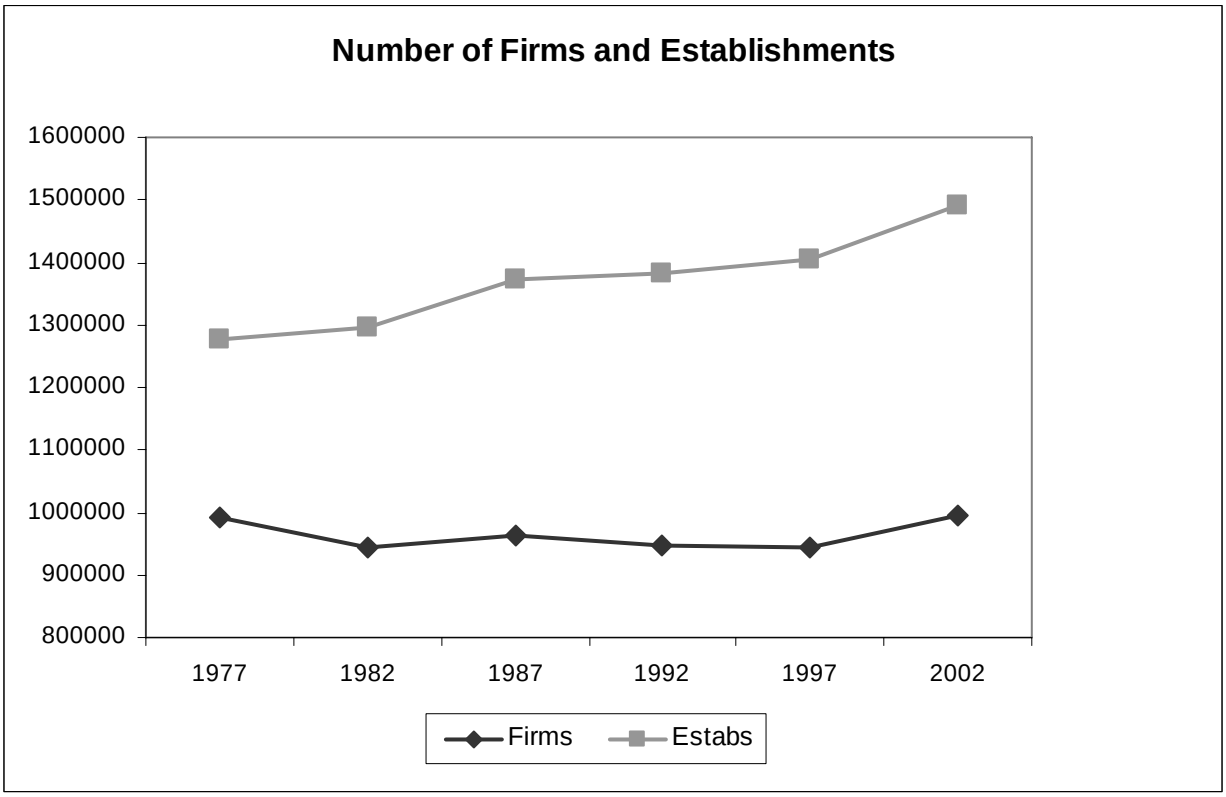
Note: Cells that cannot be disclosed are denoted by "D."

Table 4: Transition Matrix by Age of Oldest Establishment in the Firm

Cohort	Birth Share	Tenure over Census Years 0	5	10	15	20	25
Employment							
1977	93.4	14.1	11.2	9.6	4.8	17.7	36.1
1982	3.3	0.3	2.3	0.7	D	D	
1987	1.8	0.7	1.1	D	D		
1992	0.9	0.4	0.5	D			
1997	0.4	0.4	D				
2002	0.2	0.2					
Real Sales							
1977	93.3	13.8	8.4	9.0	3.7	22.3	36.0
1982	4.0	0.3	3.5	0.3	D	D	
1987	1.4	0.8	0.6	D	D		
1992	1.0	0.6	0.4	D			
1997	0.2	0.2	D				
2002	0.2	0.2					

Note: Cells that cannot be disclosed are denoted by “D.”

Figure 1: Summary Statistics for Retail Trade Sample, 1977-2002



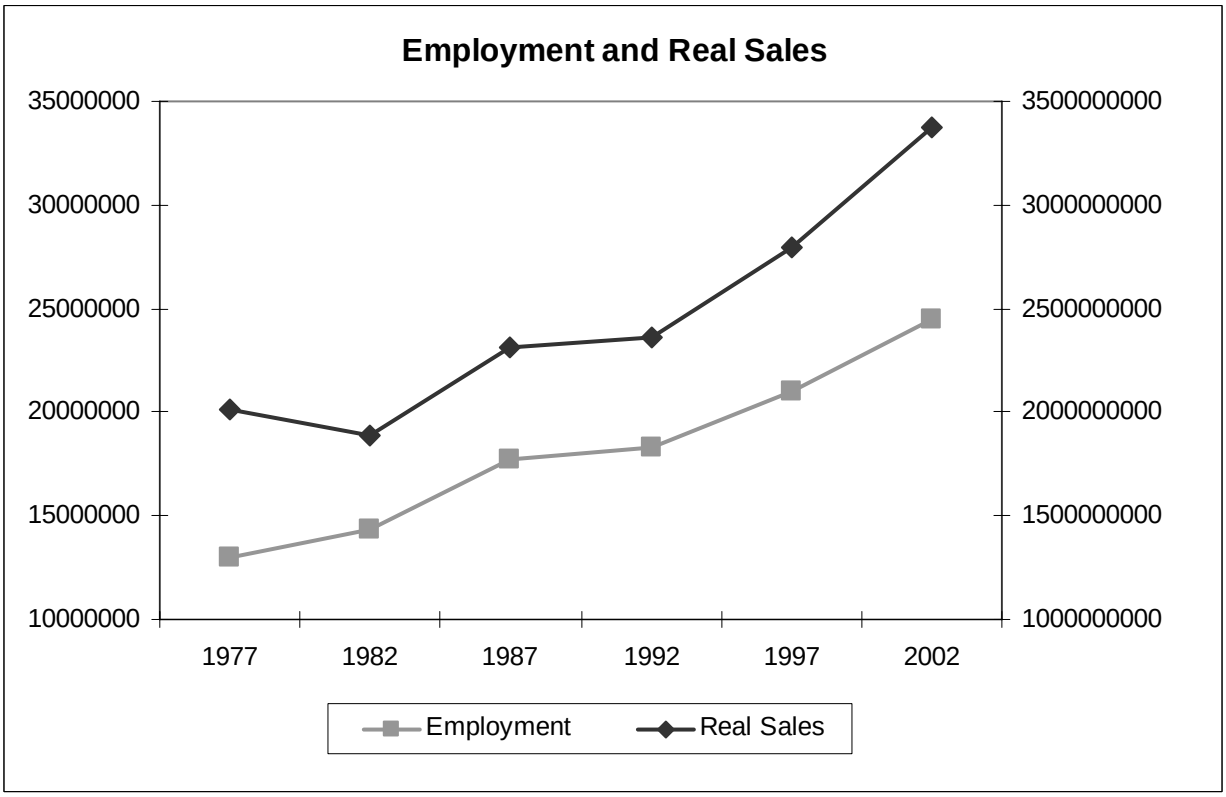
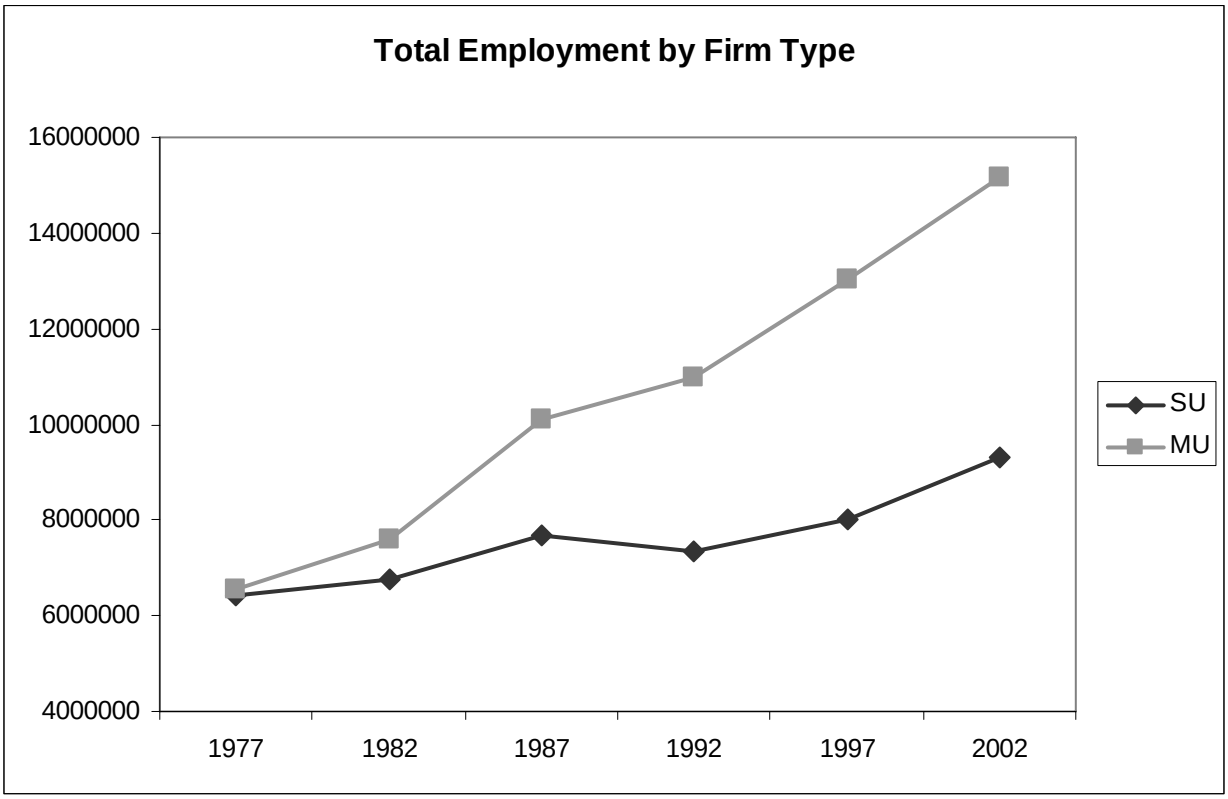


Figure 2: Total Employment and Real Sales by Firm Type



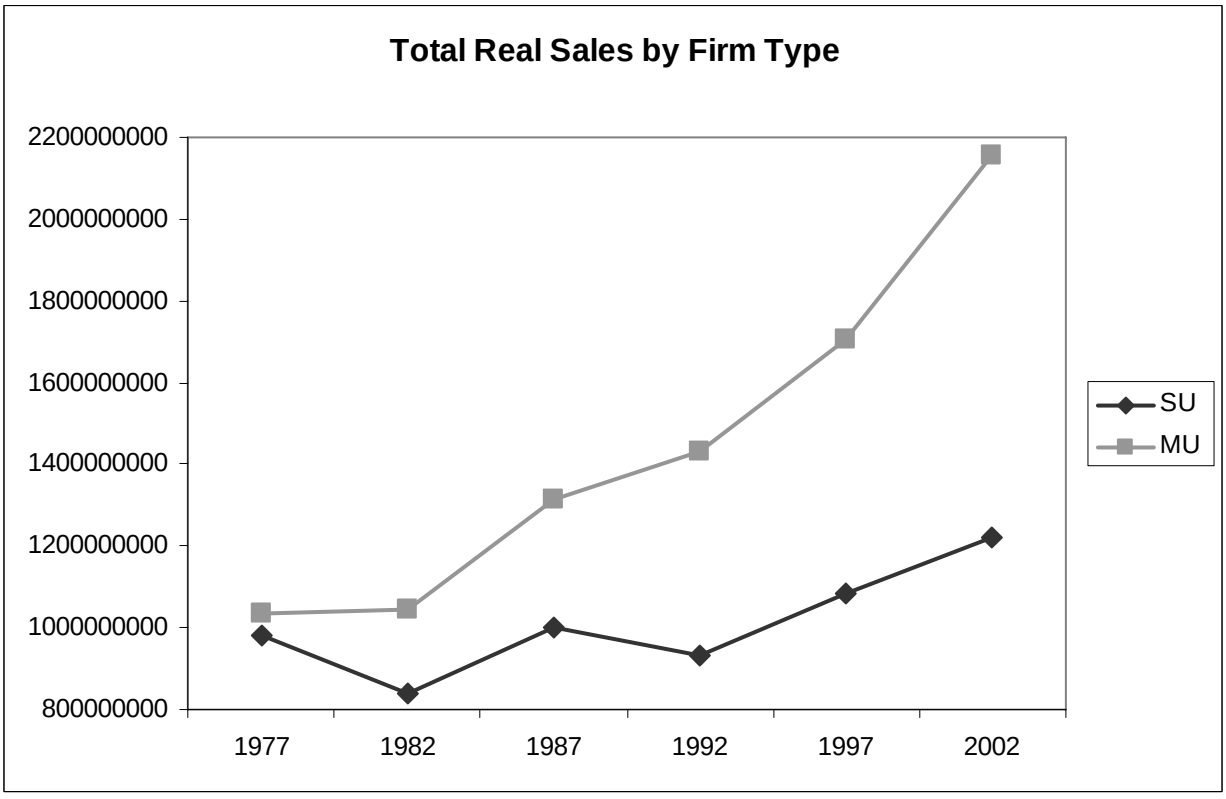
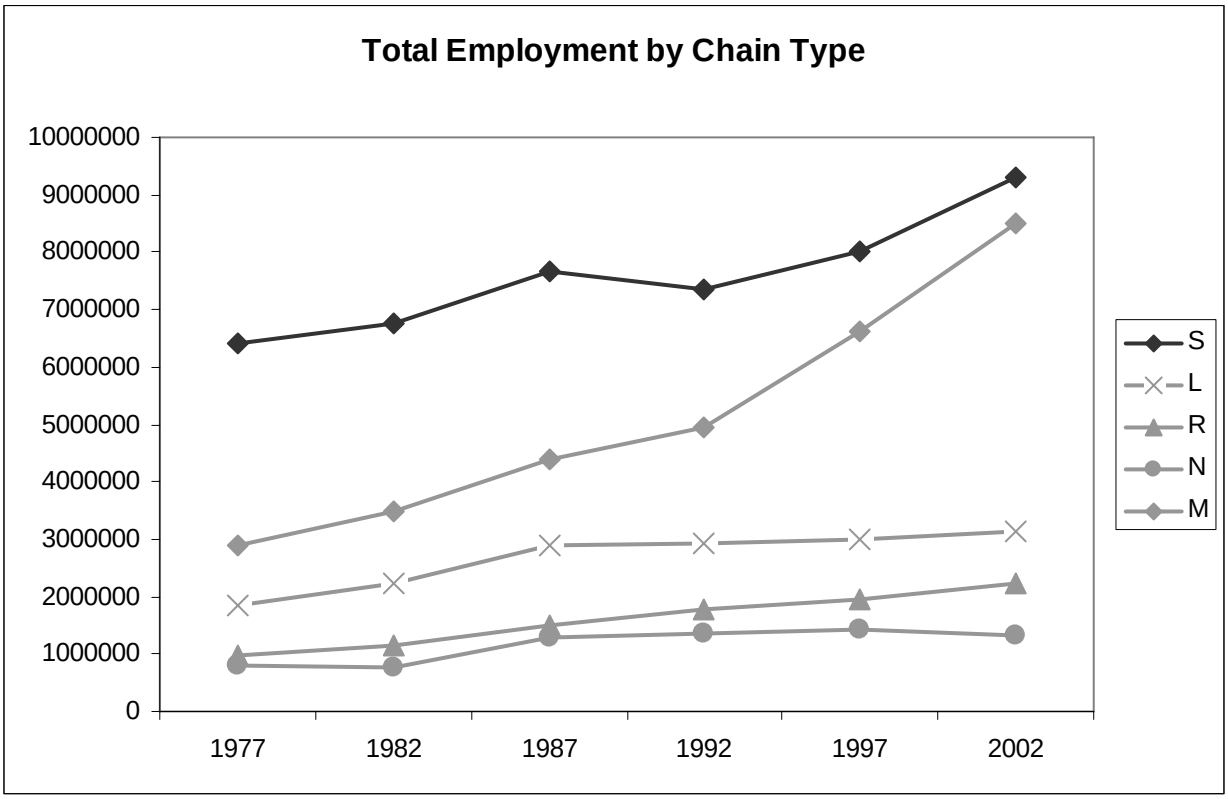


Figure 3: Total Employment and Real Sales by Chain Type



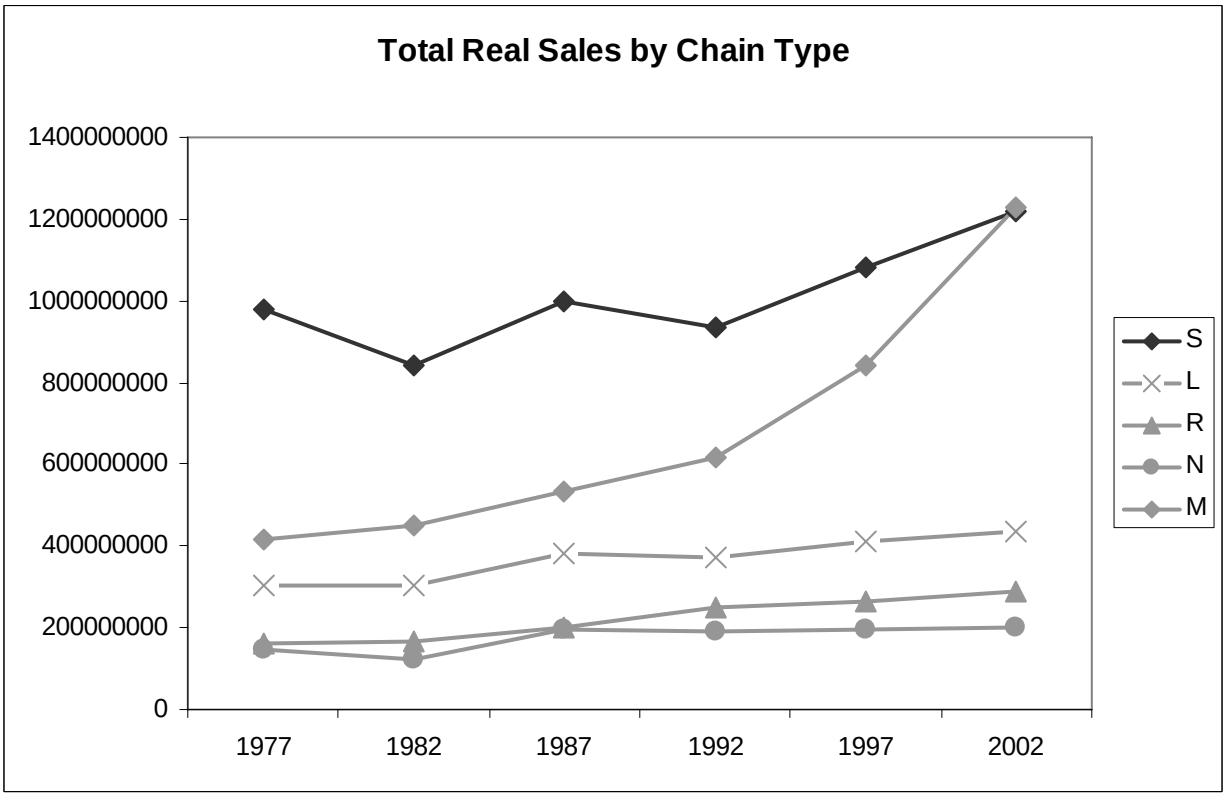


Figure 4: Total Number of Firms by Chain Type

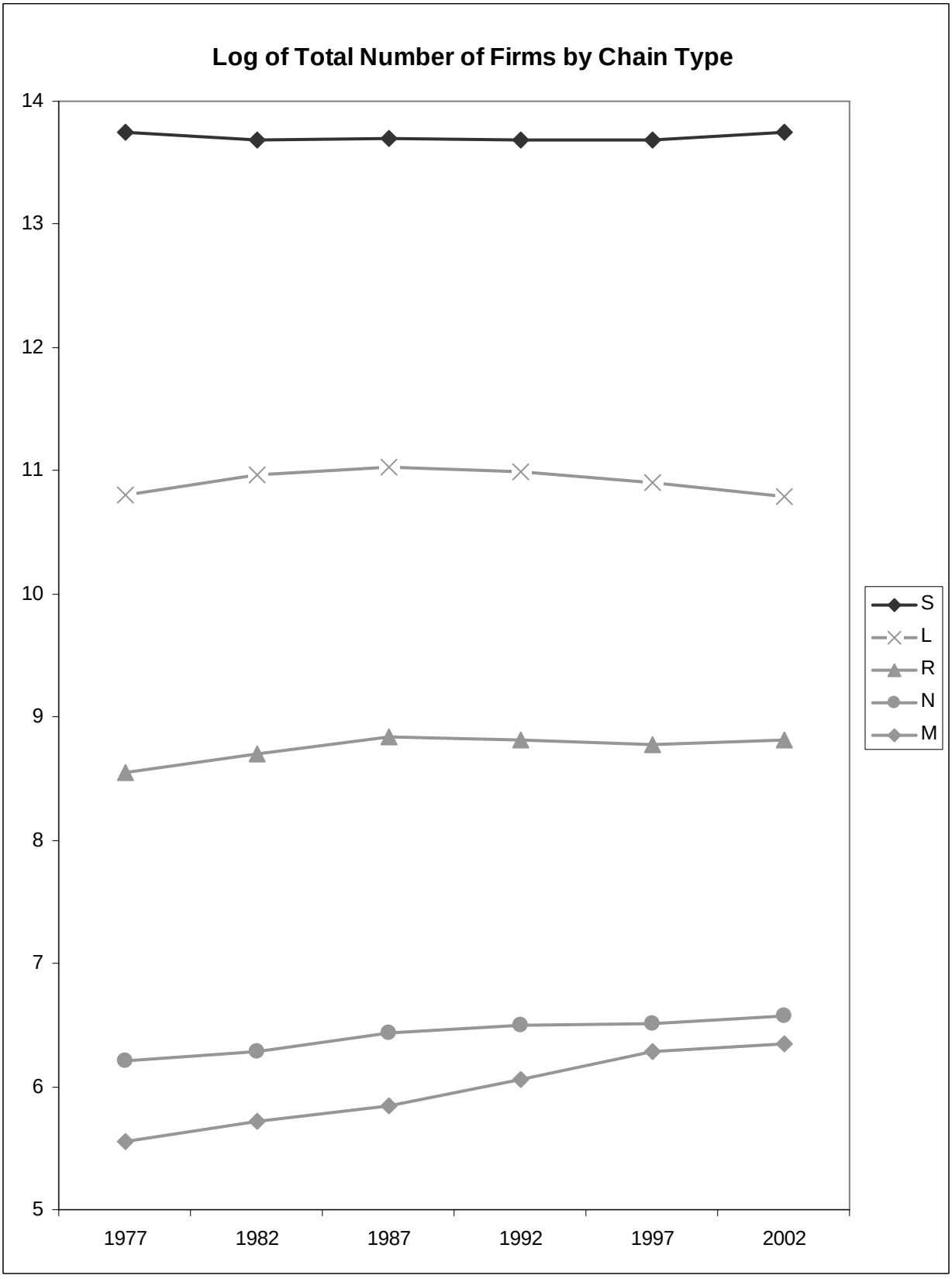


Figure 5: Total Number of Establishments by Chain Type

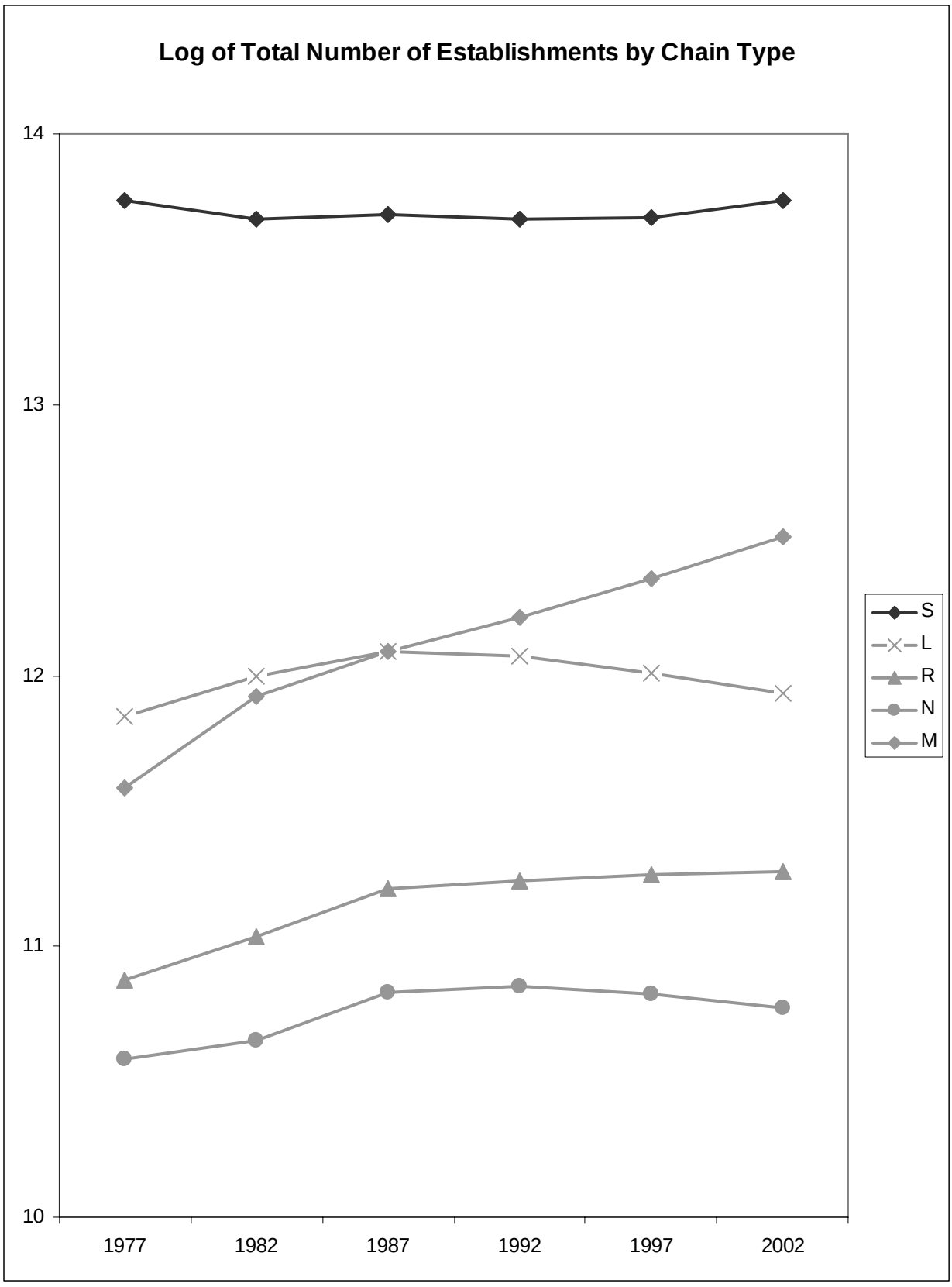


Figure 6: Mean Number of Establishments in Logs by Chain Type

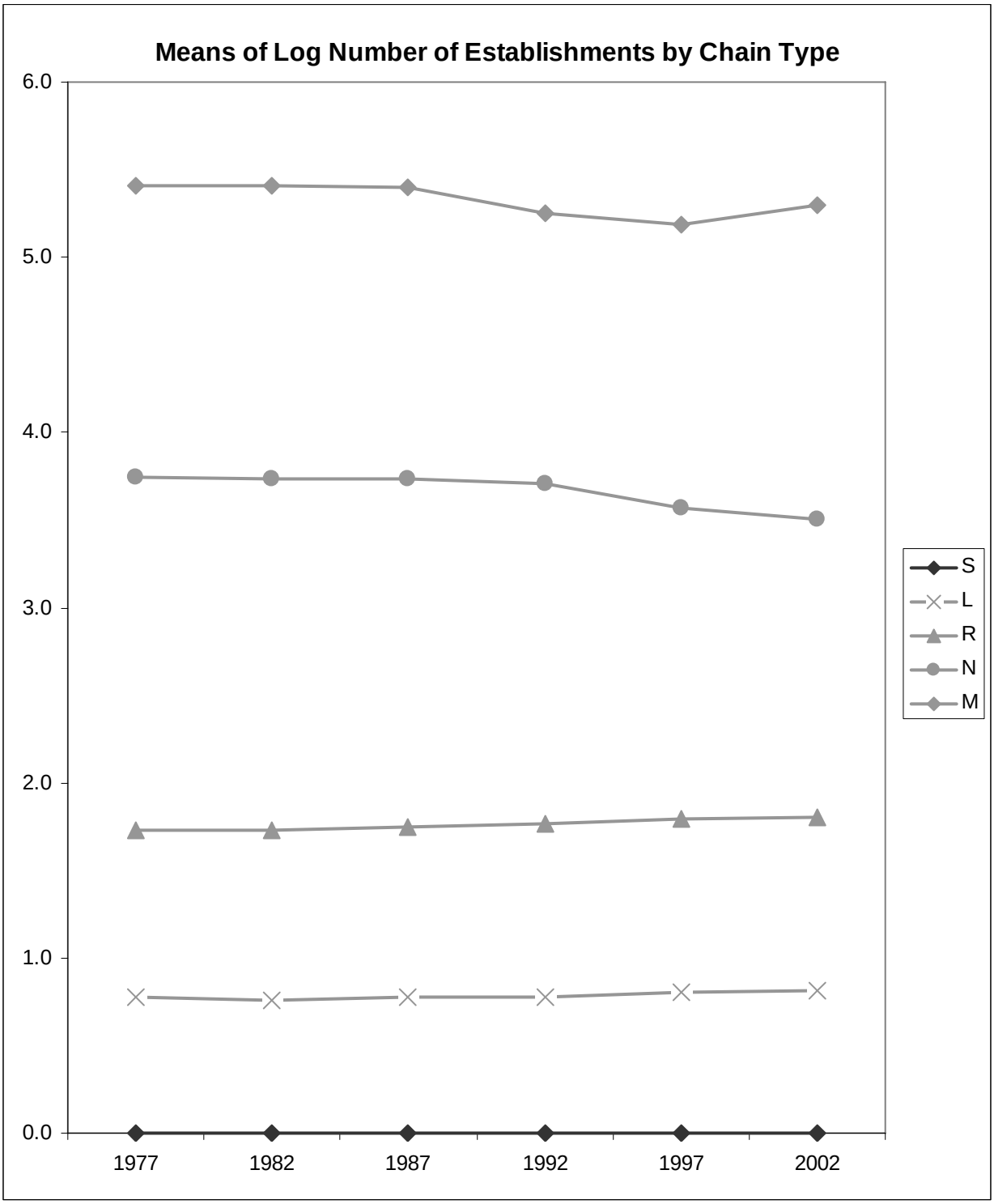


Figure 7: Mean Employment in Logs by Chain Type

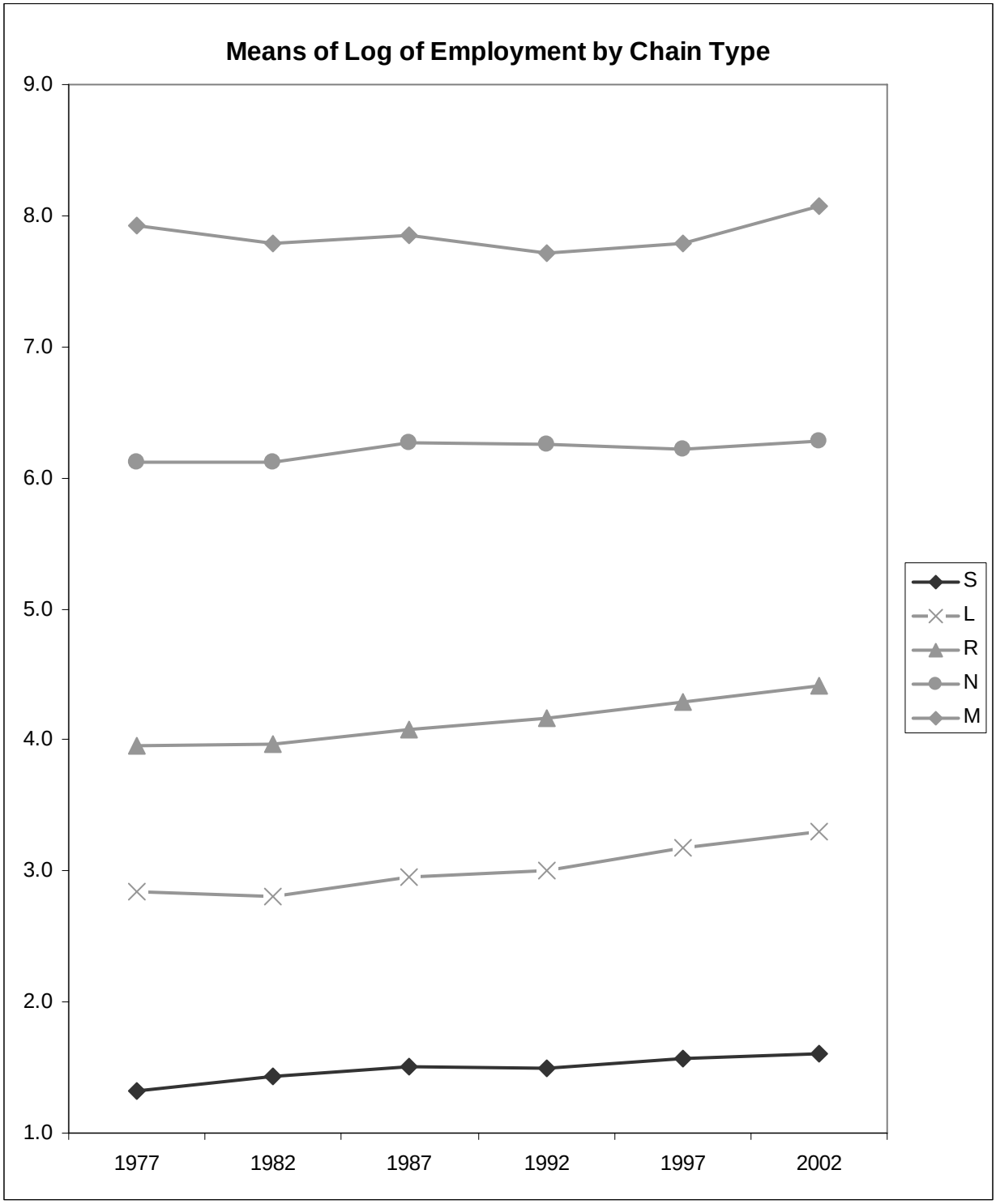


Figure 8: Mean Real Sales in Logs by Chain Type

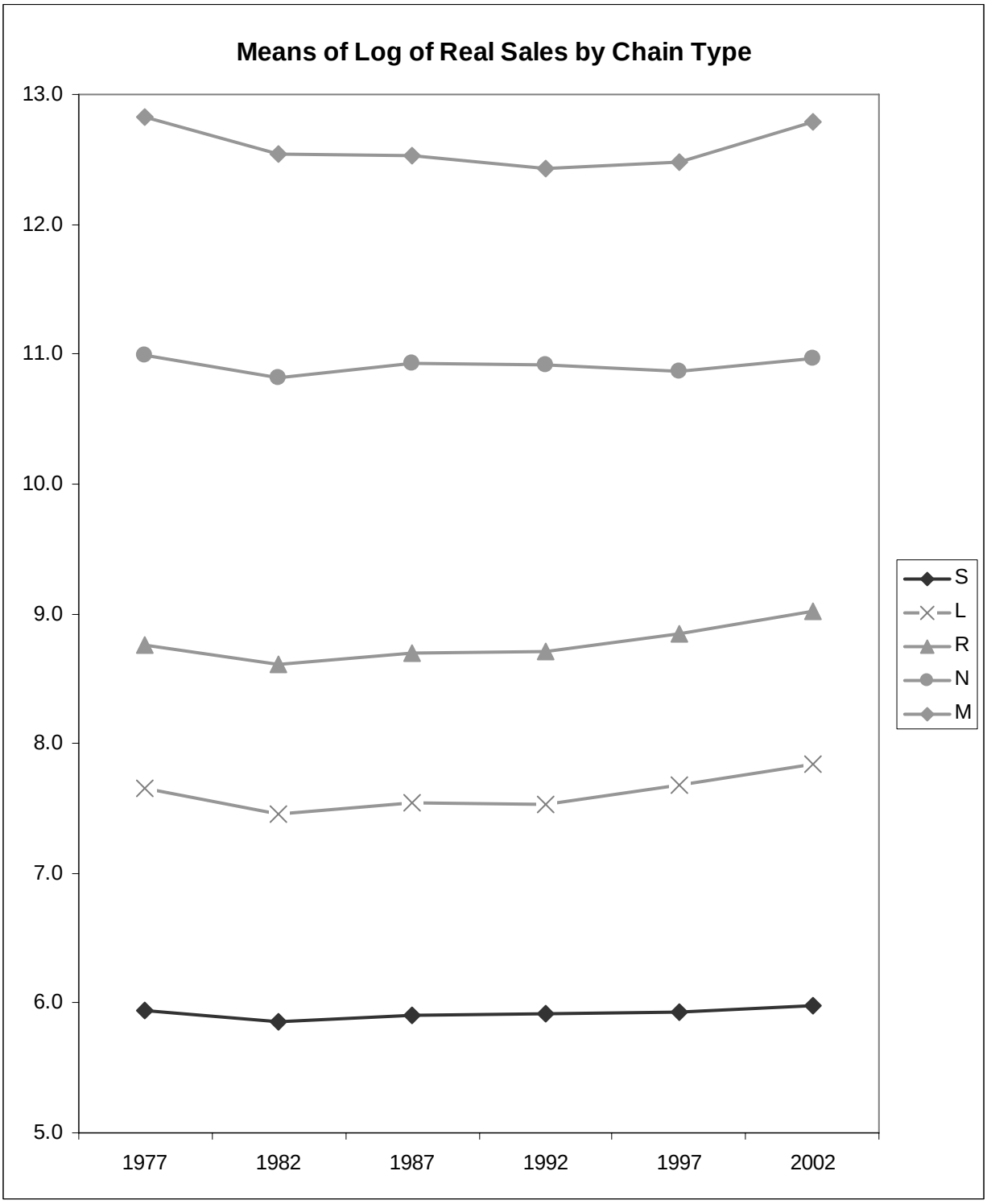


Figure 9: Cohort Analysis of Average Log Employment Per Firm for 2002 Mega Firms

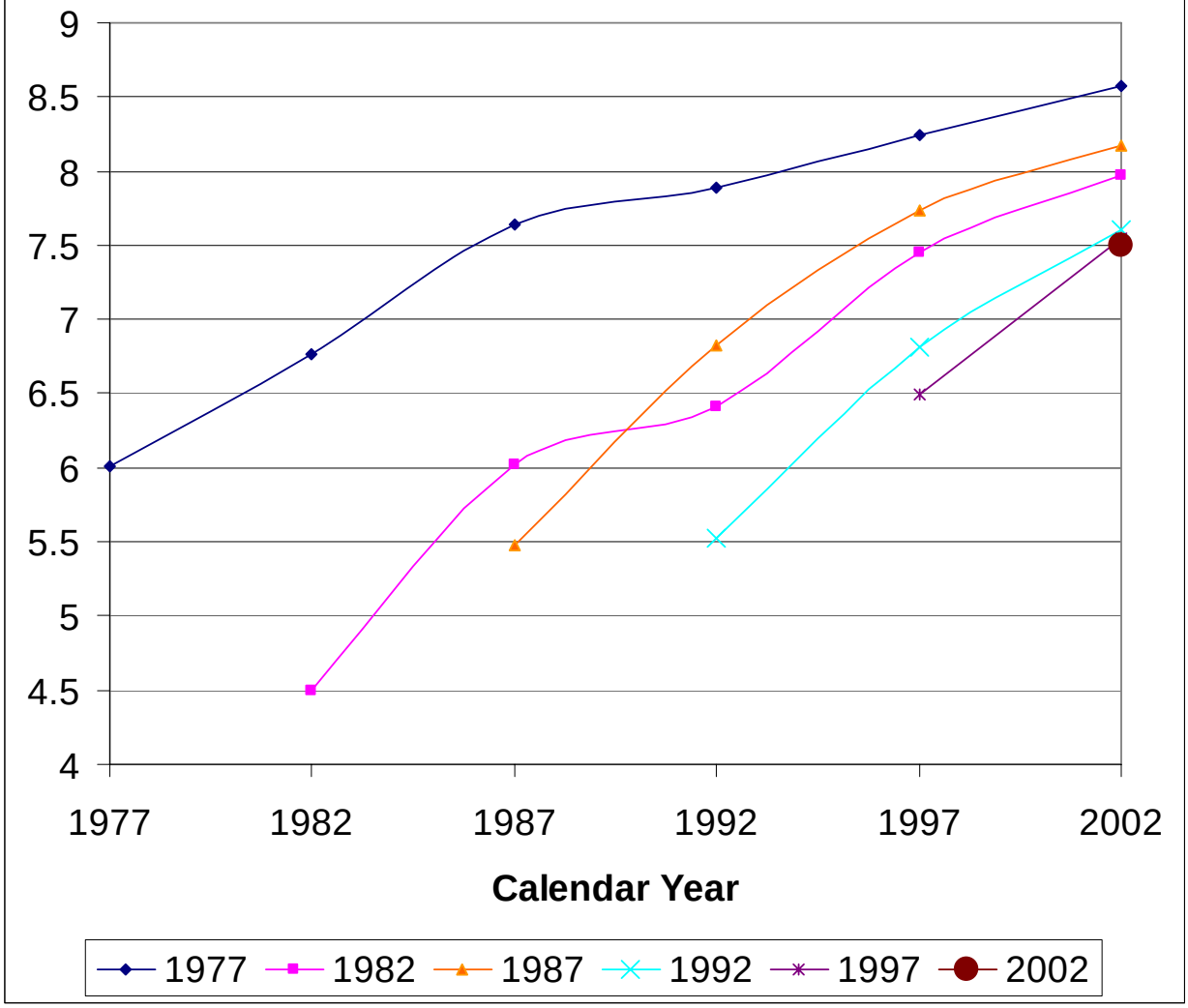


Figure 10: Cohort Analysis of Average Log Real Sales Per Firm for 2002 Mega Firms

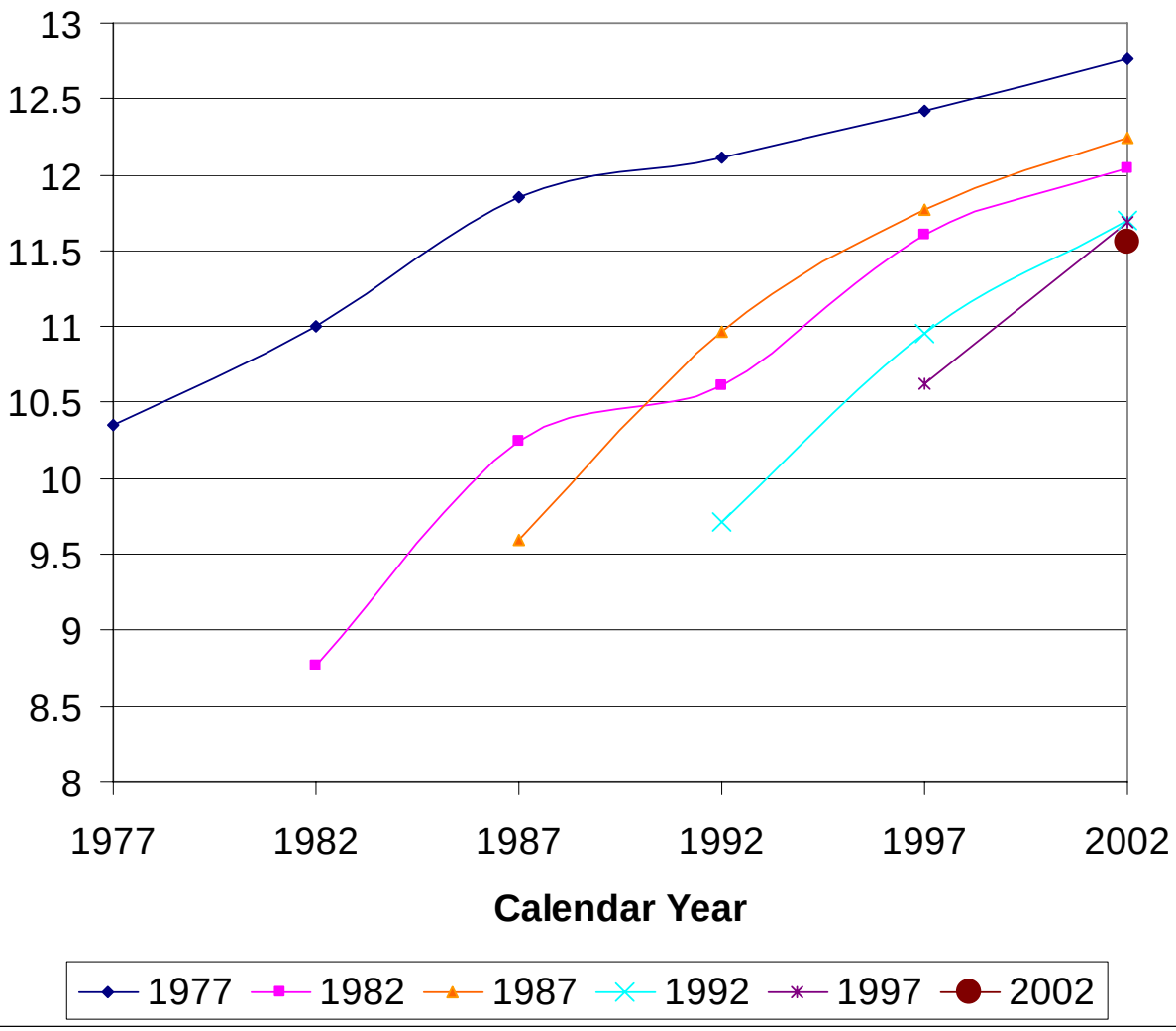


Figure 11: Cohort Analysis of Average Log Number of Establishments Per Firm for 2002 Mega Firms

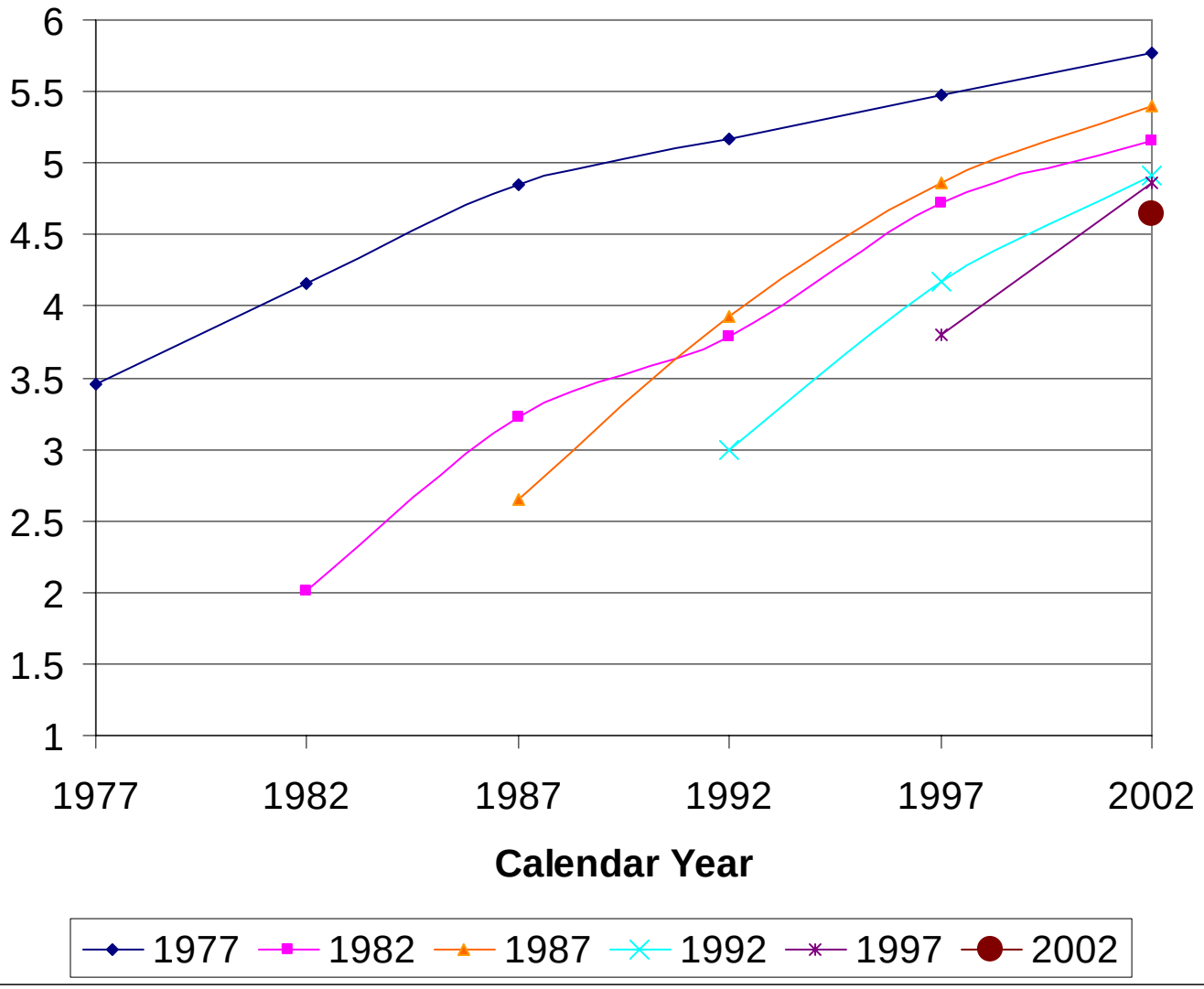


Figure 12: Cohort Analysis of Average Number of States Per Firm for 2002 Mega Firms

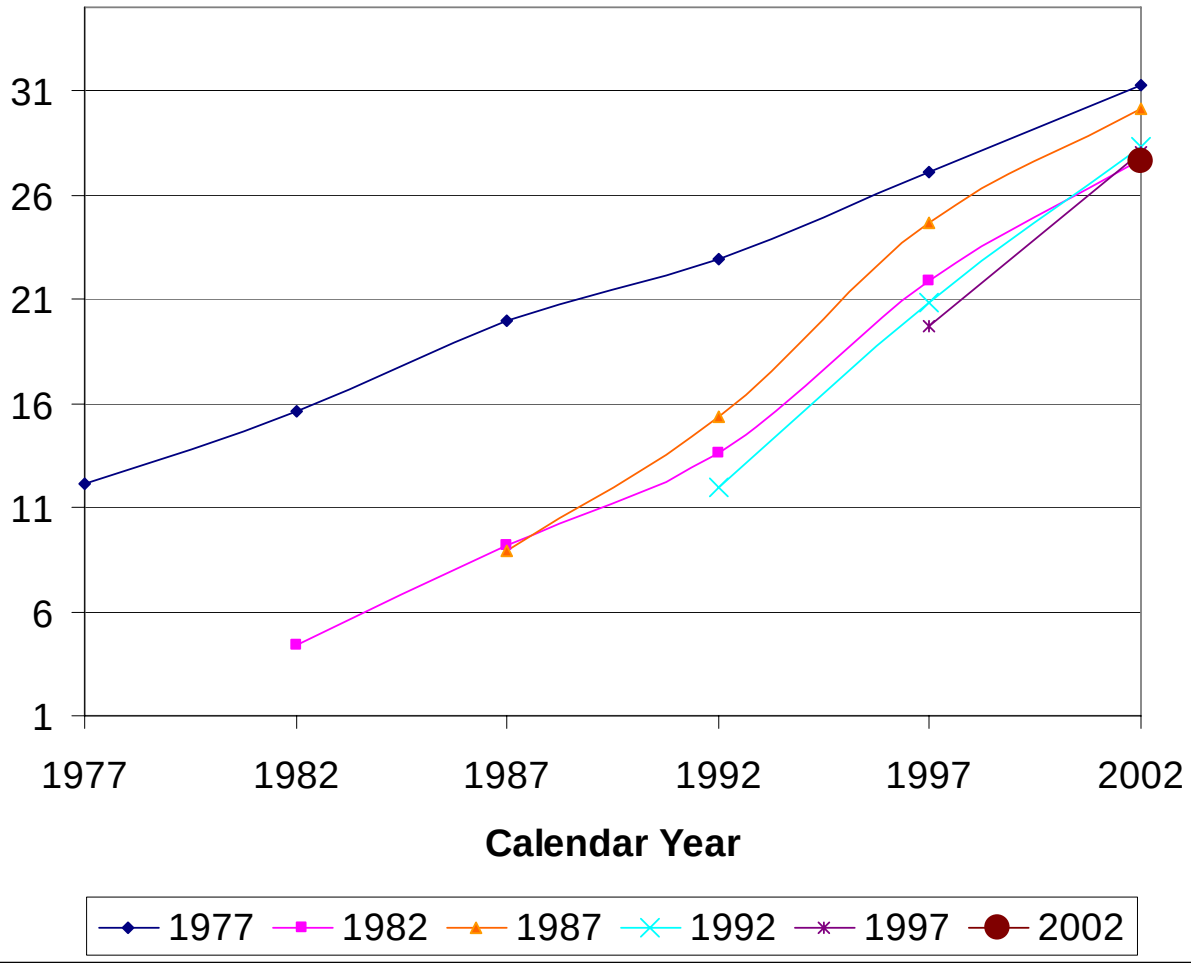


Figure 13: Cohort Analysis of Average Share of Publicly Traded Firms for 2002 Mega Firms